

融合特征金字塔与可变形卷积的高密度群养猪计数方法

王 荣^{1,2} 高荣华^{1,3} 李奇峰^{1,3} 冯 璐^{1,3} 白 强^{1,2} 马为红^{1,3}

(1.北京市农林科学院信息技术研究中心,北京 100097;
2.西北农林科技大学信息工程学院,陕西杨凌 712100;
3.国家农业信息化工程技术研究中心,北京 100097)

摘要: 针对猪只人工计数方法消耗时间和劳动力,育肥猪较为活跃且喜好聚集,图像中存在大量的高密度区域,导致猪只之间互相粘连、遮挡等问题,基于 SOLO v2 实例分割算法,提出了一种自然养殖场景下融合多尺度特征金字塔与二代可变形卷积的高密度群养猪计数模型。通过优化模型结构来减少计算资源的消耗与占用。将科大讯飞给出的猪只计数的公开数据集划分为猪只分割数据集和猪只盘点测试集,利用猪只分割数据集获得较好的分割模型,然后在猪只盘点测试集中测试盘点准确率,实现猪群分割和猪只计数。实验结果表明,本文提出的高密度猪只计数模型的分割准确率达到 96.7%,且模型内存占用量为 256 MB,为改进前的 2/3,实现了遮挡、粘连和重叠情况下的猪只个体高准确率分割。在含有 500 幅猪只图像计数测试集中,模型计算猪只数量误差为 0 时的图像数量为 207 幅,较改进前提高 26%。模型计算猪只数量误差小于 2 头猪的图像数量占测试图像总数量的 97.2%。模型计算猪只数量误差大于 3 头猪的图像数量占总体图像数量比例仅为 1%。最后,对比基于 YOLO v5 的群养猪计数方法,本文模型具有更优的分割效果和计数准确率,验证了本文方法对群养猪只计数的有效性。因此,本文模型既实现了高密度猪群的精准计数,还通过优化模型结构大大降低了模型对计算设备的依赖,使其适用于养殖场内猪群在线计数。

关键词: 高密度群养猪; 计数模型; 实例分割; SOLO v2; 多尺度特征金字塔网络; 可变形卷积

中图分类号: S431.9; TP391.4 **文献标识码:** A **文章编号:** 1000-1298(2022)10-0252-09

OSID:



High-density Pig Herd Counting Method Combined with Feature Pyramid and Deformable Convolution

WANG Rong^{1,2} GAO Ronghua^{1,3} LI Qifeng^{1,3} FENG Lu^{1,3} BAI Qiang^{1,2} MA Weihong^{1,3}

(1. Research Center of Information Technology, Beijing Academy of Agriculture and Forestry Sciences, Beijing 100097, China
2. College of Information Engineering, Northwest A&F University, Yangling, Shaanxi 712100, China
3. National Engineering Research Center for Information Technology in Agriculture, Beijing 100097, China)

Abstract: Pig counting is a critical task in large-scale breeding and intelligent management, and the manual counting method is time-consuming and labor-intensive. Since fattening pigs are more active and like to congregate, there are many high-density areas in the image, which causes problems such as adhesion and occlusion between pigs, making pig counting difficult. Based on the SOLO v2 instance segmentation algorithm, a high-density group pig counting model in natural breeding scenarios was proposed, which incorporated multi-scale feature pyramids and deformable convolutions networks version 2. Further, by optimizing the model structure, the consumption and occupation of computing resources were reduced. The pig count dataset published by iFLYTEK was divided into two parts: pig segmentation dataset and pig count test set. The pig segmentation dataset was used to train the segmentation model to achieve herd segmentation and pig counts, and the inventory accuracy was tested in the pig inventory test set. The experimental results showed that the high-density pig counting model proposed had a

收稿日期: 2022-04-17 修回日期: 2022-07-25

基金项目: 北京市自然科学基金项目(4202029)

作者简介: 王荣(1993—),女,博士生,主要从事深度学习在农业中的应用研究,E-mail: rongw@nwafu.edu.cn

通信作者: 高荣华(1976—),女,副研究员,主要从事计算机视觉在农业中的应用研究,E-mail: gaorh@nrcita.org.cn

segmentation accuracy of 96.7% and a model weight size of 256 MB, which was 1/3 less than that before the improvement. Each improved method proposed improved the model's segmentation accuracy and achieved high-accuracy segmentation of individual pigs in the case of occlusion, adhesion, and overlap. In the 500-image pig counting test set, the number of images when the model counted pigs with an error of 0 was 207, which was 26% more than that before the improvement. The number of images when the error of the model counting pigs was less than two pigs, accounted for 97.2% of the total number of test images. The number of images with a counting error of more than three pigs, accounted for only 1% of the total number of images. Finally, for the pig herd counting method based on YOLO v5, the model had a better segmentation effect and counting accuracy, proving the method's effectiveness for counting pigs in groups. Therefore, the model presented not only achieved accurate counting of high-density pig herds, but it also significantly reduced the model's reliance on computing equipment by optimizing the model structure, which made it suitable for online counting of pig herds on real farms.

Key words: high-density pig herd; counting model; instance segmentation; SOLO v2; multi-scale feature pyramid network; deformable convolutional

0 引言

随着智慧农业和信息化养殖技术的发展,大规模集成化养殖进程明显加快,视觉与人工智能技术在养殖中的应用逐渐成为新的发展趋势^[1]。生猪在不同生长阶段转栏时需要到舍内猪只重新盘点计数,现有人工巡检方式费时费力,还会因为猪只跑动造成计数不准确,利用信息化技术精确检测生猪群体数量能够提升养殖效率。但舍内猪只数量较多,粘连、遮挡等客观因素给群体盘点计数带来较大挑战,因此,迫切需要一种适合高密度养殖模式下的非接触式生猪群体精准盘点计数方法,实现高效智慧养殖^[2-3]。

传统非接触式盘点计数方法以传统图像处理方法为主。首先利用光学方法确定目标猪只典型颜色,再结合形态学和区域生长方法统计生猪群体数量^[4]。除颜色外,研究者还利用目标猪只纹理特征,实现视觉范围内猪只数量远程统计^[5]。然而这些方法均需要人工选择猪只颜色和纹理等特征,在高密度养殖模式下造成特征提取不充分,使得群体计数准确率降低。

随着深度学习与计算机视觉技术发展,卷积神经网络(Convolutional neural network, CNN)表现出较强的特征提取能力^[6-7],被很多研究者应用到盘点计数领域^[8-16]。高云等^[17]通过感知多尺度特征获取预测密度图,并分析密度图积分实现猪只数量估计,改进后的模型平均绝对误差和均方根误差分别为 1.74 和 2.28。TIAN 等^[18]提出新的 CNN 学习图像特征到密度图的映射关系,将盘点误差降为 1.67。然而密度图积分估计群体数量的方法无法精准识别猪只区域,造成盘点计数误差较高。为准确获取猪只区域,研究者利用目标检测网络 YOLO (You only look once)检测目标猪只,再结合深度相

机 Kinect 对其进行分割,实现实时的猪体分割^[19]。利用 Faster R - CNN (Faster regions with CNN feature) 和神经网络结构搜索算法(Neural architecture search, NAS)检测群养猪图像中的猪只位置并识别猪只姿态^[20]。然而基于目标检测算法得到的猪只区域,对高密度的粘连猪只的检测误差较大,无法获取猪只边缘信息。为解决此问题, SARWAR 等^[21]使用全卷积网络(Fully convolution network, FCN)在无人机(UAV)捕获的航空图像中分割牲畜个体,实现了牲畜边缘信息的提取,将整个系统的召回率从 90% 提高到 98%。XU 等^[22]利用 Mask R - CNN (Mask regions with CNN feature) 分割无人机拍摄的牛羊图像并进行盘点计数,准确率达 92%。同时利用特征金字塔和非极大值抑制等方法优化 Mask R - CNN 并将其应用于猪只盘点,取得较好的盘点效果,但对高密度养殖模式下生猪群体盘点计数准确率只有 86%^[23]。上述研究成果证明实例分割算法可获取猪只边缘信息,对稀疏生猪盘点计数的误差较小,但在猪只互相粘连、严重遮挡等高密度养殖模式下,猪只边缘分割精度仍有待提高,影响盘点计数精度。

WANG 等^[24]提出的 SOLO v2 (Segmenting objects by locations)可以精准快速地分割高密度下实例物体,成为主流的实例分割网络。因此本文以 SOLO v2 为基础,融合多尺度特征金字塔网络 (Feature pyramid network, FPN) 和二代可变形卷积^[25] (Deformable convolutional networks version 2, DCN v2),提出一种高密度养殖模式下的生猪群体盘点计数模型,解决猪只之间粘连、遮挡等问题,并根据分割结果计算单个猪圈内猪只数量,同时优化模型结构,降低其对计算资源的依赖,为生猪群体盘点计数在线监测提供技术支撑。

1 计数模型构建

1.1 数据来源

本文试验数据采用科大讯飞公开的生猪盘点数据集,如表 1 所示。数据采集时间为 2019 年 5 月 1 日到 6 月 30 日,分别在两个群养的育肥猪舍中安装俯视摄像头,选取早中晚差异较大的图像,每幅图像包括 18~31 头数量不等的猪只,共 700 幅,图像分辨率为 2 048 像素×1 563 像素和 1 920 像素×1 080 像素,保存为 .jpg 格式,24 位彩色图像。本文利用 200 幅图像作为猪只分割数据集,并按照比例 8:2 将其划分为训练集和验证集,然后利用剩余 500 幅图像作为猪只盘点测试集测试模型盘点生猪数量的效果,将猪只分割数据集中 50% 的图像进行随机翻转,增加数据集的多样性。

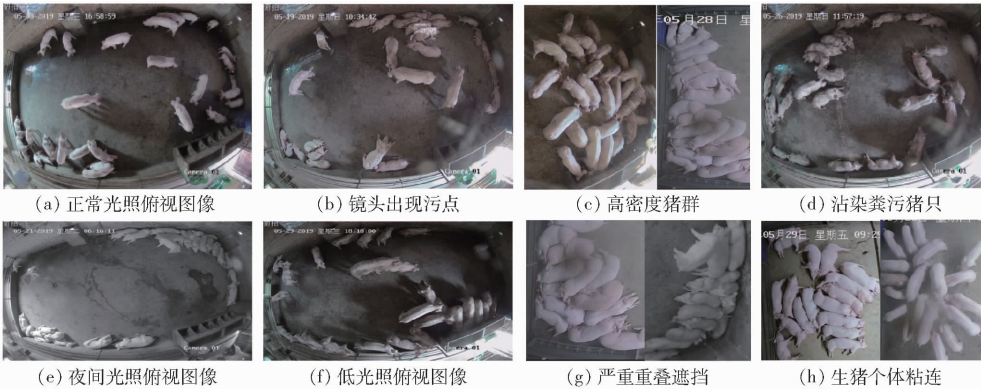


图 1 自然场景下群养猪俯视图像

Fig. 1 Top view images of group pigs in natural scene

1.2 生猪群体盘点计数模型

1.2.1 模型网络结构

为适应高密度养殖模型下的多姿态猪只盘点计数,以 ResNet 为基础建立盘点计数骨干网络,利用猪只特征金字塔提取多尺度盘点特征图,再整合特征图分批输入至猪只实例分割模块,根据分割结果

表 1 生猪盘点数据集				
Tab. 1 Analysis of pig inventory data set				
数据集	图像数量/幅	增强方法	图像包含猪只数量/头	图像分辨率/(像素×像素)
猪只分割数据集	200	随机翻转	25~31	2 048×1 563、1 920×1 080
猪只盘点测试集	500		18~29	2 048×1 563、1 920×1 080

图 1 为群养猪俯视图像,每幅图像猪只数量在 18~31 头之间,由于育肥猪较为活跃,群养时容易聚集,因此每幅图像中均存在高密度区域。加之猪场环境脏乱,拍摄的盘点图像受环境和猪群姿态影响较大,这些高密度区域出现较多的模糊、脏乱、严重遮挡、重叠和粘连等情况,因此增加了生猪个体分割的难度,也为盘点计数带来挑战。

实现生猪群体盘点计数,模型结构如图 2 所示。盘点数据集被输入到高密度生猪群体盘点计数模型后,被划分为 $S \times S$ 网格,经过骨干网络和猪只特征金字塔得到不同层次特征图,输入到猪只实例分割模块的不同支路进行多任务学习,最终输出不同位置的语义类别和实例掩膜。利用矩阵非极大值

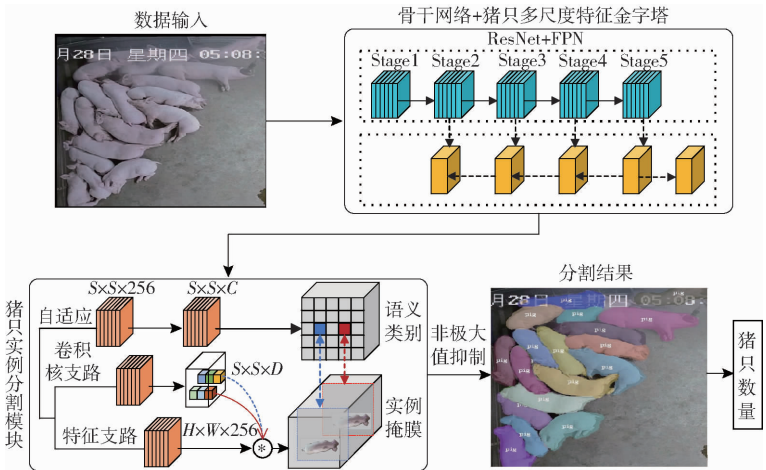


图 2 高密度生猪群体盘点计数模型结构

Fig. 2 Structure of high-density pig population inventory count model

抑制算法 (Matrix non-maximum suppression, Matrix NMS) 去除得分较低的实例掩膜, 输出最终的群养猪实例分割结果, 实现生猪群体盘点计数。为适用于粘连、密集和遮挡等情况下的群体精准计数, 本文选

取 ResNet18、ResNet34、ResNet50 和 ResNet101 共 4 种骨干网络寻找适用于生猪群体盘点的最优模型, 减少训练中的过拟合现象, 骨干网络具体结构见表 2。

表 2 骨干网络结构
Tab. 2 Backbone network structure

结构名称	ResNet18	ResNet34	ResNet50	ResNet101
Stage1	7 × 7, 64, 步长 2 3 × 3, 最大池化层, 步长 2			
Stage2	$\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$
Stage3	$\begin{bmatrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 8$
Stage4	$\begin{bmatrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1\,024 \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1\,024 \end{bmatrix} \times 36$
Stage5	$\begin{bmatrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2\,048 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2\,048 \end{bmatrix} \times 3$

1.2.2 猪只特征金字塔构建

单一视角下获取的大视野生猪盘点数据集, 往往存在不同尺度猪只个体, 因此, 融合 ResNet 骨干

网络和猪只特征金字塔提取不同层级语义信息, 解决图像中多尺度猪只分割问题, 达到提高模型分割精度的目的, 具体结构如图 3 所示, 图中括号内数值

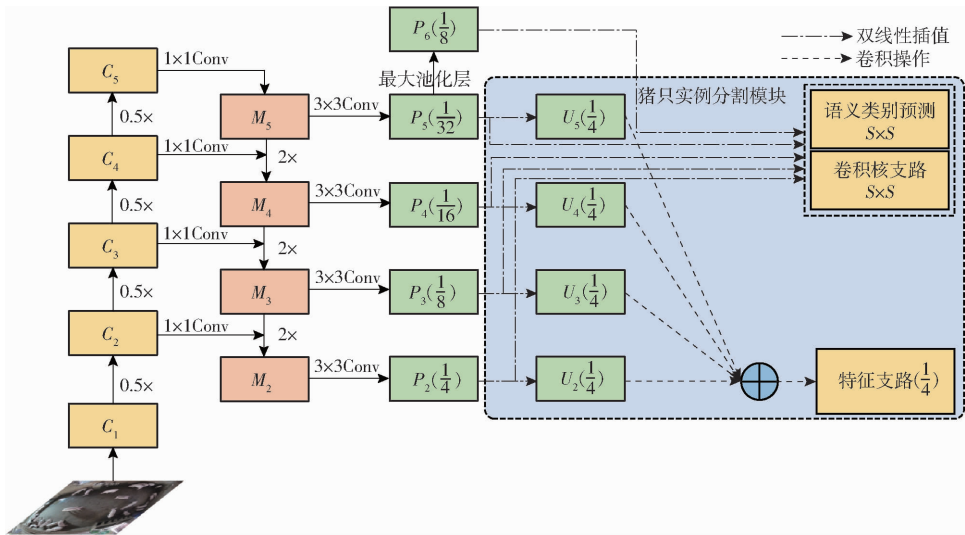


图 3 猪只特征金字塔结构图

Fig. 3 Pyramid structure diagram of pig characteristics

表示特征图尺寸占原图尺寸的比例。
高层特征包含的语义信息较多, 低层特征的特征损失较少, 因此抽取包含高层语义信息的特征图 C_5 , 经过 1×1 卷积操作得到通道数为 256 的特征图 M_5 , 抽取特征图 $C_2 \sim C_4$, 分别与上一层上采样后的特征图相加, 得到 256 通道的融合特征图 $M_2 \sim M_4$ 。 $M_2 \sim M_5$ 经过 3×3 卷积操作之后得到包含语义类别和位置信息的 4 种不同尺度特征图 $P_2 \sim P_5$, 其

中, P_5 经过最大池化层后得到 P_6 。利用双线性插值将 $P_2 \sim P_6$ 缩放至 $S \times S$ 大小 (其中 S 为 12、16、24、36、40), 作为卷积核支路和语义类别预测支路的输入。利用重复卷积和双线性插值将 $U_2 \sim U_5$ 缩放至相同尺度后进行特征融合, 输入至特征支路预测猪只掩膜。

1.2.3 猪只实例分割模块

猪只实例分割模块包含语义类别预测和实例掩

膜预测两个分支,猪只特征金字塔输出的特征图 F 被输入到实例掩膜预测分支后,利用特征支路和卷积核支路获取最终的群养猪掩膜预测结果 M ,计算公式为

$$M_{i,j} = G_{i,j} * F \tag{1}$$

式中 $M_{i,j}$ ——群养猪的掩膜预测结果

$G_{i,j}$ ——中心位置为 (i,j) 的掩膜内核

实例掩膜预测支路有两个分支,因此需要两个输入 F 和两个掩膜内核 G 生成猪只掩膜 M 。被输入到卷积核支路的特征图 F 尺寸为 $S \times S \times 256$,经过 4 个 3×3 卷积和一个 $3 \times 3 \times D$ 的操作后生成最终掩膜内核 $G_{i,j}$,对于每个网格,卷积核分支预测输出维度为 D ,当卷积操作为 1×1 时, D 与 F 的通道数相同,卷积操作为 3×3 时, D 为 F 通道数的 9 倍。

特征支路利用重复的 3×3 卷积操作、组归一化^[26]、ReLU 和上采样等操作,将猪只特征金字塔输出的特征图 $P_2 \sim P_5$ 进行特征融合,融合之后的特征图再次经过 1×1 卷积操作、组归一化以及 ReLU 激活函数得到最终的掩膜特征图。

1.2.4 卷积结构优化

传统卷积使用固定大小的卷积核作为滑动窗口对特征图 F 进行采样,然后点乘 ω 求和,卷积公式为

$$y(P_0) = \sum_{P_n \in R_k} \omega(P_n) \cdot F(P_0 + P_n) \tag{2}$$

式中 $y(P_0)$ ——输出的卷积结果

R_k ——感受野的大小和扩张,取值为 $\{(-1, -1), (-1, 0), \dots, (0, 1), (1, 1)\}$

P_n ——以 P_0 为中心卷积核范围内的所有采样点

$\omega(P_n)$ ——卷积核的权重矩阵

$F(P_0 + P_n)$ ——以 P_0 为中心卷积核范围内的图像特征

卷积操作中固定的感受野对多尺度目标的鲁棒性较差,易造成较大的特征损失。为进一步提高模型对多尺度猪只的鲁棒性,本文引入第 2 代的可变形卷积(Deformable convolutional,DCN v2)。在传统卷积感知输入特征时加入调整偏移量,调整来自不同空间位置和单元的特征调制振幅。当调制振幅为 0 时,该位置的特征信号不被接收,实现模型对空间位置特征的自主学习和选择。 $y(P_0)$ 计算公式为

$$y(P_0) = \sum_{P_n \in R_k} \omega(P_n) \cdot F(P_0 + P_n + \Delta P_n) \Delta m_n \tag{3}$$

式中 ΔP_n 、 Δm_n ——调整偏移量

仅加入调整偏移量 ΔP_n 后得到可变形卷积 DCN^[27],加入 Δm_n 后得到 DCN v2,提高了模型对群养猪图像的特征提取能力和对多尺度猪只的鲁

棒性。

1.2.5 模型结构优化

为减少模型对计算资源的消耗,对模型冗余通道剪枝,实现较轻量猪只实例分割模块。将模块支路中的 4 个卷积块减少为 3 个,生成的掩膜图通道维数由 512 降为 256,同时将尺度范围修改为 $[(1, 64), (32, 128), (64, 256), (128, 512), (256, 2\,048)]$,降低冗余的模型参数量。输入较大尺寸的图像虽然能提高模型准确率,但仍会降低模型推理速度,增加模型对计算资源的消耗,因此,在保证模型准确率基础上,本文降低多尺度训练和测试时输入图像分辨率,将多尺度训练时图像分辨率设为 $[(852 \text{ 像素}, 512 \text{ 像素}), (852 \text{ 像素}, 480 \text{ 像素}), (852 \text{ 像素}, 448 \text{ 像素}), (852 \text{ 像素}, 416 \text{ 像素}), (852 \text{ 像素}, 384 \text{ 像素}), (852 \text{ 像素}, 352 \text{ 像素})]$,测试时由 $(1\,333 \text{ 像素}, 800 \text{ 像素})$ 降为 $(852 \text{ 像素}, 512 \text{ 像素})$,实现对生猪群体盘点计数模型结构优化。

1.2.6 损失函数和评价指标

生猪群体盘点计数模型的损失函数由 Focal 语义类别预测损失函数 L_{cate} 和 Dice 掩膜预测损失函数 L_{mask} 组成,利用猪只个体分类误差和猪只掩膜预测误差,共同判定模型收敛程度。

采用平均精度(Average precision,AP)评价模型性能,使用交并比(Intersection over union,IoU)在 0.50~0.95 的联合阈值上取平均值,并利用 IoU 取 0.50 或 0.75 时的 AP 来判断模型精度,评价指标值越高,证明模型分割猪只精度越高。

2 实验结果与分析

2.1 实验设置

实验采用 1 块 GeForce GTX 2080Ti 型 GPU 进行训练,显存为 11 GB,基于 Windows 10 操作系统和 Pytorch 框架搭建深度学习算法网络训练平台,其中 Python 版本为 3.7.1,Pytorch 版本为 1.10.0,Keras 版本为 2.2.4,CUDA API 版本为 10.2,cuDNN 版本为 8.0.5。

2.2 基础模型选择

ResNet 骨干网络包含多种不同层数的特征提取网络,使用相同的生猪盘点数据集和多尺度训练方法训练具有不同骨干网络的 SOLO v2 模型,获得适用于生猪群体盘点计数的 ResNet 骨干网络,实验结果如表 3 所示。

由表 3 可知,基于 ResNet50 和 ResNet101 的 SOLO v2 模型效果最好。在 IoU 为 0.50~0.95、0.75 时基于 ResNet101 的 SOLO v2 的 AP 最高,IoU 为 0.50 时 SOLO v2 + ResNet101 比 SOLO v2 +

表 3 不同骨干网络的 SOLO v2 模型猪只分割结果

Tab.3 Pig segmentation results of SOLO v2 model with different backbone networks

模型	AP/%			内存占用 量/MB
	IoU 为	IoU 为	IoU 为	
	0.50 ~ 0.95	0.50	0.75	
SOLO v2 + ResNet18	57.2	91.7	69.8	252
SOLO v2 + ResNet34	53.4	89.7	62.8	329
SOLO v2 + ResNet50	58.4	94.9	72.0	352
SOLO v2 + ResNet101	59.6	93.8	74.2	497

ResNet50 的 AP 低 1.1 个百分点,且内存占用量比其高 145 MB。为确保模型轻量化和精确性,本文选择 ResNet50 作为骨干网络改进 SOLO v2,使其适用于生猪盘点在线计数。

2.3 模型训练结果

模型训练过程中采用梯度下降法 (Stochastic gradient descent,SGD) 优化模型权重,设置初始学习率为 0.01,动态衰减因子为 0.9,正则化权值衰减系数为 0.000 1,在 150 轮和 250 轮时降低学习率,总训练轮数为 400 轮,训练时每次输入 GPU 的图像数量为 10。利用多尺度分辨率图像训练本文模型,训练结果如图 4 所示。

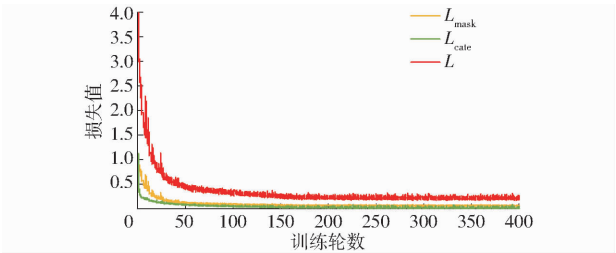


图 4 高密度生猪群体盘点计数模型的训练结果

Fig.4 Training results of high-density pig population inventory counting model

由图 4 可知,随着训练轮数的增加,3 个损失函数总体呈下降趋势,趋于稳定后,在固定区间内波动。其中掩膜预测损失函数在 0.06 ~ 0.08 区间波动,语义类别预测损失函数在 0.01 ~ 0.02 区间波动,总体损失函数在 0.2 ~ 0.25 区间震荡,掩膜损失函数收敛区间略高于语义类别预测损失函数收敛区间,其原因是掩膜损失函数计算依赖于猪只分割区域像素值矩阵,计算得到的损失函数值偏高。因此总体收敛趋势表明本文模型分割猪只效果较好。

2.4 不同模型之间测试结果对比

为验证本文模型对高密度生猪群体的分割精度,设计了多组消融对比实验。使用相同骨干网络、训练方式、训练集和测试集,对比 SOLO v2、模型结构优化后的 SOLO v2、引入 DCN 的结构优化 SOLO v2 及本文模型的猪只分割效果,SOLO v2 优化模型

的测试结果对比如表 4 所示。

表 4 不同模型的测试结果对比

Tab.4 Comparison of test results between different models

模型	AP/%			内存占 用量/ MB	每幅图 像推理 时间/s
	IoU 为	IoU 为	IoU 为		
	0.50 ~ 0.95	0.50	0.75		
SOLO v2	58.4	94.9	72.0	352	0.86
结构优化 SOLO v2	55.5	93.3	65.3	237	0.67
结构优化 SOLO v2 + DCN	57.8	94.4	71.0	251	0.67
本文模型	63.3	96.7	80.3	256	0.68

以 ResNet50 为基础骨干网络的 SOLO v2 模型在 IoU 为 0.50 时分割的准确率达到了 94.9%,保存的内存占用量为 352 MB,对 SOLO v2 结构进行优化,模型内存占用量为 237 MB,减少 115 MB,证明模型结构优化方法的有效性。但结构优化后的模型在 IoU 为 0.50 时准确率下降 1.6 个百分点,IoU 为 0.50 ~ 0.95 和 IoU 为 0.75 时的 AP 也出现一定程度下降,因此,分别引入 DCN 和 DCN v2,提升模型准确率。在结构优化后的 SOLO v2 模型中引入 DCN,AP (IoU 为 0.50) 提高 1.1 个百分点,IoU 为 0.50 ~ 0.95 和 0.75 时 AP 分别提高了 2.3、5.7 个百分点。而本文模型进行结构优化并引入 DCN v2 后,相较于仅优化模型结构的 SOLO v2 的 AP (IoU 为 0.50) 提高 3.4 个百分点,IoU 为 0.50 ~ 0.95 和 IoU 为 0.75 时的 AP 分别提高 7.8、15 个百分点,因此,本文模型提高了生猪群体分割准确率,且模型内存占用量为 256 MB,较 SOLO v2 模型降低 1/3,适用于真实养殖场内生猪群体的在线盘点计数。

表 4 证明了本文模型对生猪群体分割准确率最高,鲁棒性最好。因此,对不同模型分割生猪群体结果进行可视化分析,如图 5 所示,图像中重叠、粘连的区域使用红色虚线框标出。由图 5 可知,SOLO v2 模型对稀疏猪群分割效果较好,红色虚线标记的猪只粘连、堆叠区域,SOLO v2 优化模型存在将粘连猪只视为同一头猪或将一头猪只分割为多头的错分现象。而本文模型对严重重叠、粘连区域的分割错误较少。因此,图 5 进一步验证了本文模型对重叠、粘连生猪群体有较好的分割效果,同时,本文模型推理每幅图像时间仅为 0.68 s,能够实现生猪群体实时盘点。

2.5 不同模型生猪盘点效果对比分析

为进一步验证本文模型泛化性,采用 500 幅带有高密度区域的猪群盘点测试数据集,如图 6 所示,

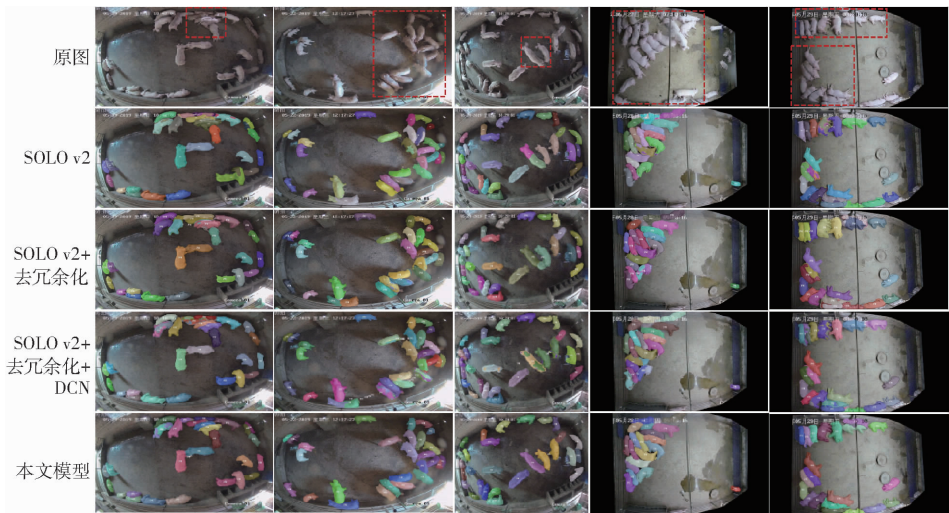


图5 不同模型分割生猪个体的可视化结果

Fig.5 Visual analysis of results of segmentation of individual pigs by different models

500 幅测试图像中每幅图像的猪只数量在 18 ~ 29 头不等。利用 SOLO v2 及不同优化模型盘点图像生猪数量,分析不同模型群体盘点计数结果,如表 5 所示。

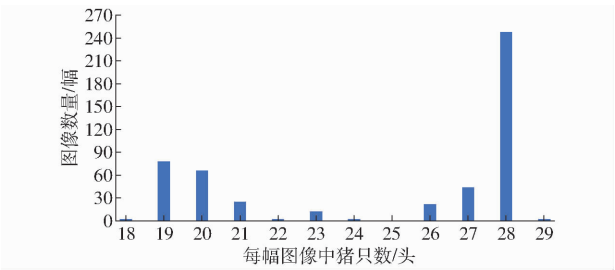


图6 500 幅盘点测试图像中的猪只数量分布

Fig.6 Distribution of pig numbers in 500 inventory test images

表 5 不同模型群体盘点计数结果对比
Tab.5 Comparison of population count results of different models

模型	误差为					幅
	0 头猪	1 头猪	2 头猪	3 头猪	3 头猪	
SOLO v2	164	192	108	30	6	
结构优化 SOLO v2	86	171	141	65	37	
结构优化 SOLO v2 + DCN	143	211	106	35	5	
本文模型	207	201	78	9	5	

实验过程中,设置得分阈值为 0.5,由表 5 可知,本文模型盘点误差为 0 头猪的图像数量为 207 幅,远高于其他模型,相较于改进前的 SOLO v2,盘点误差为 0 头猪的图像数量增加了 26%。其中,盘点计数误差为 0 或 1 头猪的图像数量占总体测试图像的 81.6%,而 SOLO v2 仅占 71.2%;误差为 2 头猪以内的图像数量占整体测试图像的 97.2%,误差大于 3 头猪的图像仅 5 幅,占整体测试图像的 1%,

相较于表 5 中的 SOLO v2 优化模型,本文模型实现了高密度养殖模式下的高精度盘点计数。

2.6 分析与讨论

目前基于深度学习的猪只计数研究中常使用目标检测算法检测图像中的猪只个体,然后根据检测结果计算猪只数量。这种方法对于稀疏离散分布的猪只图像中检测效果较好,当图像中存在高密度区域时,猪只粘连、遮挡、重叠等问题使得目标检测模型准确率降低,而基于实例分割算法的猪只计数可减少高密度猪群的干扰。因此,本文复现了目前性能较好的 YOLO v5 模型对群养猪进行计数的准确率,并和本文基于实例分割模型的猪只计数算法进行对比,结果如表 6 所示。

表 6 不同方法对猪群计数准确率对比结果
Tab.6 Comparison results of pig herd counting by different methods

模型	误差为				%
	0 头猪	2 头猪	3 头猪	4 头猪	
YOLO v5	27.2	63.0	86.4	94.8	5.2
本文模型	41.4	81.6	97.2	99.0	1.0

对 YOLO v5 模型进行 200 轮训练后,IoU 为 0.50 时,模型目标检测的准确率达到 99.2%,在猪只盘点测试集测试猪只计数效果,不同方法对猪群计数的对比结果如表 6 所示。本文模型计数正确的猪群图像数量约为 YOLO v5 的两倍,且本文模型计数误差小于 3 头猪时准确率为 97.2%,比 YOLO v5 模型提高 10.8 个百分点;而 YOLO v5 模型计数误差大于 3 头猪的图像数量是本文模型的 5 倍。综上所述,YOLO v5 作为目前性能最优的目标检测模型,经训练之后,虽然检测准确率较高,但其应用在猪只计数时准确率远低于本文模型,验证了本文选择实

例分割算法用于猪只计数优于基于目标检测的猪只计数方法。

为进一步验证基于实例分割的猪只计数效果优于基于目标检测算法的猪只计数方法,对猪只计数测试集的分割结果和检测结果进行可视化分析,如图 7 所示。对于稀疏分散的群养猪图像,图像分割和检测均可以较为准确地找到猪只位置,但当猪体形状不是规则的矩形时,目标检测方法会出现误检

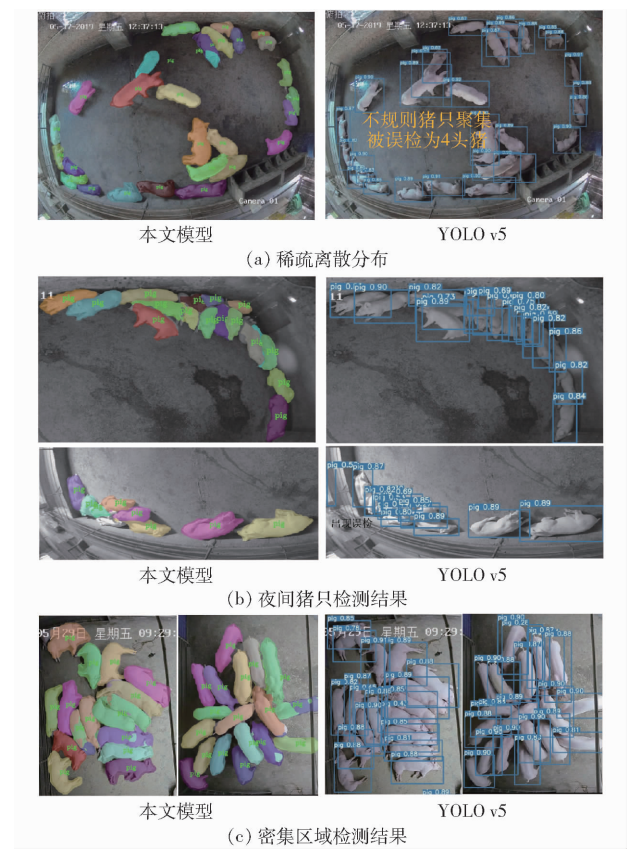


图 7 不同方法检测猪只可视化效果

Fig. 7 Visualization of pig detection by different methods

的情况,而本文模型没有出现,如图 7a 所示。夜间群养猪图像为红外图像,对本文模型的影响较小,而 YOLO v5 会将背景检测为猪只个体,导致计数不准确,如图 7b 所示。对于群养猪图像的高密度区域,猪只粘连、重叠现象较多,本文模型误分割的现象相对较少,而由于猪体不是规则的矩形形状,且互相粘连,YOLO v5 的检测框中常常包含多头猪,不同猪只之间相互影响,造成 YOLO v5 对高密度区域内的猪群计数准确率较低。综上,育肥猪较为活跃,喜好群居,在猪只计数测试集中出现了大量的高密度区域,而本文模型的准确率远高于 YOLO v5 模型,可视化效果也优于目标检测结果,因此,本文提出的基于实例分割模型的猪只计数方法更适用于高密度下的群养猪盘点计数。

3 结束语

本文模型利用猪只特征金字塔和 DCN v2 提高对多尺度目标的特征提取能力和分割能力,同时优化模型结构,减少对计算资源的消耗和占用,并与目前使用最多的目标检测模型 YOLO v5 进行了对比。实验结果表明,本文模型在高密度养殖模式下分割生猪群体准确率达到 96.7%,且本文模型内存占用量较 SOLO v2 减少 96 MB,降低 1/3,达到较好的猪群分割效果。盘点计数预测误差为 0 的图像数量为 207 幅,较 SOLO v2 增加 26%,其中盘点计数误差为 0 头猪或 1 头猪的图像数量占整体测试图像的 81.6%,较 SOLO v2 提高 10.4 个百分点,本文盘点误差在 2 头猪以内的图像数量占比为 97.2%,误差大于 3 头猪的图像仅占 1%。且本文模型计数正确的猪群图像数量约为 YOLO v5 的 1.52 倍,计数误差大于 3 头猪的图像数量是基于 YOLO v5 模型的 1/5。

参 考 文 献

[1] 周道雷,李保明,施正香.我国养猪产业生产现状及其发展方向[J].农业工程学报,2005,19(增刊1):101-105.
ZHOU Daolei, LI Baoming, SHI Zhengxiang. Current status and development strategy of swine industry in China[J]. Transactions of the CSAE, 2005, 19(Supp.1): 101-105. (in Chinese)

[2] 邹远炳,孙龙清,李玥,等.基于分布式流式计算的生猪养殖视频监测分析系统[J].农业机械学报,2017,48(增刊):365-373.
ZOU Yuanbing, SUN Longqing, LI Yue, et al. Video monitoring and analysis system for pig breeding based on distributed flow computing[J]. Transactions of the Chinese Society for Agricultural Machinery, 2017, 48(Supp.): 365-373. (in Chinese)

[3] MARCHANT J A, SCHOFIELD C P, WHITE R P. Pig growth and conformation monitoring using image analysis[J]. Animal Science, 2016, 68(1): 141-150.

[4] 张天昊,梁炎森,何志毅.图像识别计数在储备生猪统计的应用[J].计算机应用与软件,2016,33(12):173-178.
ZHANG Tianhao, LIANG Yansen, HE Zhiyi. Application of image recognition counting in pig reserve statistics[J]. Computer Applications and Software, 2016, 33(12): 173-178. (in Chinese)

[5] 梁炎森,张天昊,何志毅.畜牧养殖场图像远程采集与目标计数系统[J].桂林电子科技大学学报,2017,37(6):10-14.
LIANG Yansen, ZHANG Tianhao, HE Zhiyi. Remote image acquisition and object counting system for livestock farms[J]. Journal of Guilin University of Electronic Science and Technology, 2017, 37(6): 10-14. (in Chinese)

[6] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. Imagenet classification with deep convolutional neural networks[C]//

- Advances in Neural Information Processing Systems, 2012: 1097 – 1105.
- [7] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016: 770 – 778.
- [8] CHEN S W, SHIVAKUMAR S S, DCUNHA S, et al. Counting apples and oranges with deep learning: a data-driven approach [J]. IEEE Robotics and Automation Letters, 2017, 2(2): 781 – 788.
- [9] MADS D, HENRIK K, HENRIK S M. Plant species classification using deep convolutional neural network [J]. Biosystems Engineering, 2016, 151: 72 – 80.
- [10] GEMERT J C V, VERSCHOOR C R, METTES P, et al. Nature conservation drones for automatic localization and counting of animals [C] // Proceedings of the European Conference on Computer Vision, 2015: 255 – 270.
- [11] KAMILARIS A, PRENAFETA-BOLDU F X. Deep learning in agriculture: a survey [J]. Computers and Electronics in Agriculture, 2018, 147: 70 – 90.
- [12] PICON A, ALVAREZ-GILA A, SEITZ M, et al. Deep convolutional neural networks for mobile capture device-based crop disease classification in the wild [J]. Computers and Electronics in Agriculture, 2019, 161(1): 280 – 290.
- [13] RAHNEMOONFAR M, SHEPPARD C. Deep count: fruit counting based on deep simulated learning [J]. Sensors, 2017, 17(4): 905.
- [14] 梁勇,邱荣洲,李志鹏,等. 基于 YOLO v5 和多源数据集的水稻主要害虫识别方法 [J]. 农业机械学报, 2022, 53(7): 250 – 258.
LIANG Yong, QIU Rongzhou, LI Zhipeng, et al. Identification method of major rice pests based on YOLO v5 and multi-source datasets [J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(7): 250 – 258. (in Chinese)
- [15] SHEN Y F, ZHOU H L, LI J T, et al. Detection of stored-grain insects using deep learning [J]. Computers and Electronics in Agriculture, 2018, 145: 319 – 325.
- [16] 薛卫,程润华,康亚龙,等. 基于 GC – Cascade R – CNN 的梨叶病斑计数方法 [J]. 农业机械学报, 2022, 53(5): 237 – 245.
XUE Wei, CHENG Runhua, KANG Yalong, et al. Pear leaf disease spot counting method based on GC – Cascade R – CNN [J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(5): 237 – 245. (in Chinese)
- [17] 高云,李静,余梅,等. 基于多尺度感知的高密度猪只计数网络研究 [J]. 农业机械学报, 2021, 52(9): 172 – 178.
GAO Yun, LI Jing, YU Mei, et al. High-density pig counting net based on multi-scale aware [J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(9): 172 – 178. (in Chinese)
- [18] TIAN M X, GUO H, CHEN H, et al. Automated pig counting using deep learning [J]. Computers and Electronics in Agriculture, 2019, 163: 1 – 10.
- [19] MISO J, YOUNG C, JIHYUN S. A Kinect-based segmentation of touching-pigs for real-time monitoring [J]. Sensors, 2018, 18(6): 1746 – 1770.
- [20] RIEKERT M, KLEIN A, ADRION F, et al. Automatically detecting pig position and posture by 2D camera imaging and deep learning [J]. Computers and Electronics in Agriculture, 2020, 174: 105391.
- [21] SARWAR F, GRIFFIN A, REHMAN S U, et al. Detecting sheep in UAV images [J]. Computers and Electronics in Agriculture, 2021, 187(3): 106219.
- [22] XU B, WANG W, FALZON G, et al. Livestock classification and counting in quadcopter aerial images using Mask R – CNN [J]. International Journal of Remote Sensing, 2020(7): 1 – 22.
- [23] 胡云鸽,苍岩,乔玉龙. 基于改进实例分割算法的智能猪只盘点系统设计 [J]. 农业工程学报, 2020, 36(19): 177 – 183.
HU Yunge, CANG Yan, QIAO Yulong. Design of intelligent pig counting system based on improved instance segmentation algorithm [J]. Transactions of the CSAE, 2020, 36(19): 177 – 183. (in Chinese)
- [24] WANG X, ZHANG R, KONG T, et al. SOLO v2: dynamic and fast instance segmentation [C] // Advances in Neural Information Processing Systems, 2020: 17721 – 17732.
- [25] ZHU X, HU H, LIN S, et al. Deformable convnets v2: more deformable, better results [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2019: 9308 – 9316.
- [26] WU Y X, HE K M. Group normalization [J]. International Journal of Computer Vision, 2020, 128: 742 – 755.
- [27] DAI J F, QI H Z, XIONG Y W, et al. Deformable convolutional networks [C] // The IEEE Conference on Computer Vision and Pattern Recognition, 2017: 764 – 773.